Population Genetics Math

Introduction

This is a math-intensive lesson focusing on frequencies—how often a genotype or an allele occurs in a population—and provides a method to predict what the frequency should be for a given disorder. We're talking **population-level** genetics, considering what is likely to be found in a large population of organisms that are freely mating with each other. The central concept is the **Hardy-Weinberg equation**, which allows us to predict the **genotype frequency** (the combination of two alleles) and the **allele frequency** (the presence of a gene variant).

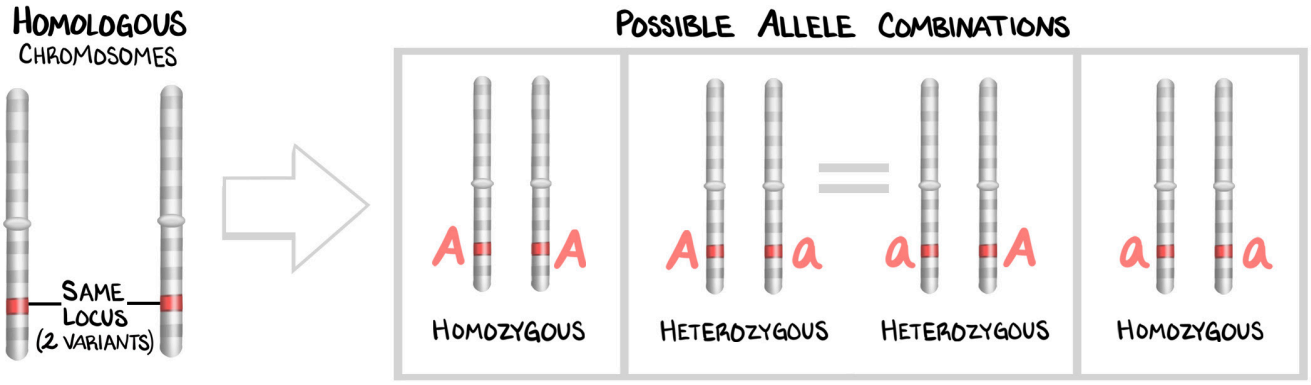

Figure 4.1: Two Alleles, Three Genotypes
A gene is the inheritable unit of code. It codes for protein. The variation of a gene is called an allele. The genotype is the combination of two alleles. The image on the left represents the two possible gene variations, the two possible alleles. The "products" on the right are NOT the possibilities of the pairings of these two chromosomes. They're all the possible pairings of gametes in all people who have two alleles for the gene in question, and there are two possible states for every gene: A or a.

Every human has **one genotype** determined by **two alleles.** To allow us to use language instead of math, we're going to consider patients to be **normal** (homozygous non-mutant), **affected** (homozygous mutant), or **carrier** (non-affected heterozygous). We're also going to assume (because it's appropriate for almost all medical genetics) that there are **only two alleles possible**—p (normal) and q (diseased). And the language "normal, affected, carrier" indicates that we will assume every condition is recessive unless specifically called out as something else.

Putting the possible combinations of alleles into a Punnett square gives the possible combinations. This matches an assessment of a population using a Southern blot. The control arm shows that we're detecting for two alleles, p and q. A band means present. The control arm ensures that the test is working. Genotype 1 (GI) has only p and equates to "normal" in the Punnett square. Genotype 3 (GIII) has a band only at q, meaning it has only q alleles, representing "affected" in the Punnett square. Genotype 2 (GII) has both p and q, representing "carrier" in the Punnett square. But there are two squares with pq. Three genotypes, four possible outcomes.

What we just did was describe the extant genotypes illustrated by alleles. We didn't say, "frequency." What just happened is that **all possible genotypes were described in a box**. That's it.

© OnlineMedEd. http://www.onlinemeded.org

Hardy-Weinberg Equation: Genotype Frequencies

If we assume there's only a good (p) and a bad (q) allele, and those are the only two possibilities, then doing a Punnett square of the possible combinations determines the three possible genotypes. They are: the normal state named $\mathbf{p^2}$, the affected stated named $\mathbf{q^2}$, and the carrier state named **pq**. The "squared" really gets to people; it may be easier to grasp by remembering the three genotypes in words: normal, carrier, and affected.

This is important to do in population genetics. The goal is to match the given information with the correct representation of letters. Let's detour a little bit, then come back to this point.

Hardy and Weinberg had a breakthrough. If all the possible allele combinations / genotypes were contained in this Punnett square, and the frequencies must sum to 1 (or 100%), then the sum of the contents of the four quadrants of the Punnett sqaure must equal 1.

$p^2 + 2pq + q^2 = 1$

The frequency of the normal state (genotype 1) is p^2, the frequency of the affected state (genotype 3) is q^2, and the frequency of the carrier state is 2pq. ALL of them must add up to a total of 1. The **genotype frequency** of the normal state, if p is known, is determined by multiplying p by itself. If the allele frequency of p is 0.6 (60%), then the genotype frequency for p^2 is 0.6 x 0.6 = .36 (36%). And if the genotype frequency is .36 (36%), the square root is the allele frequency.

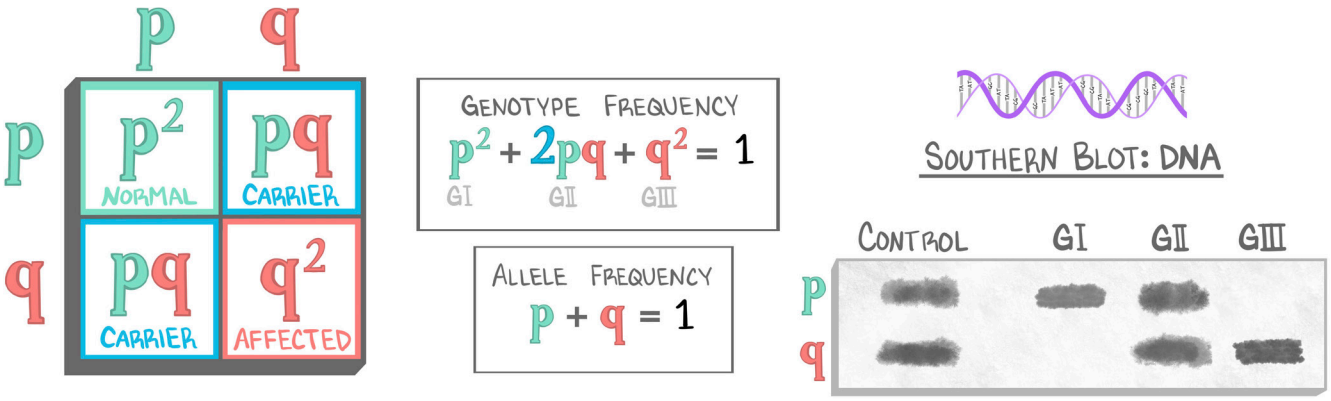

Figure 4.2: Genotypes, Phenotypes, and Frequencies
Mastering these equations and the Punnett square, or simply memorizing them and working through some problems like the ones below, can be enough to get through the exam.

Allele Frequencies

There are only two types of alleles—the good one (p) and the bad one (q). So, the frequency with which p and q occur must therefore make up the total of all alleles. So p + q = 1.

Feel what that means. Whenever p is known, so also is q (subtract p from 1). Whenever q is known, so also is p (subtract q from 1). Therefore, **just by knowing either p or q** one can deduce the predicted values for **all allele frequencies** and **all genotype frequencies** simply by doing a little math.

What do you think you're going to need to do? **Get to p or q** then plug and chug. "Getting to p or q" means recognizing the disease pattern (X-linked, autosomal recessive) and interpreting the given information. **Do not go phenotype to allele**. Do **NOT use phenotype in Hardy-Weinberg at all**. It's easier to say "normal, carrier, affected" than "genotype 1, 2, 3" or "homozygous p, etc.," but we're not using those to mean the phenotype—only the genotype.

© OnlineMedEd. http://www.onlinemeded.org

Frequencies to Quantity

Every human has **one genotype** made from **two alleles**. The **frequency** is a percentage. So, given the **population size** (the N value of a group of people) and the frequency, the number of alleles and genotypes can be derived. But be careful.

Frequencies are percentages. Each person has one genotype. T**he number of genotypes is the frequency of the genotype × the population size**.

Each person has **two alleles**. The **number of alleles is the frequency of the allele × the population size × 2**.

Allele Frequency	Size of Population	Alleles per Person	Number of Alleles
p2 = .64	10,000	1	6,400
2pq = .32	10,000	1	3,200
q2 = .04	10,000	1	400
			(total 10,000)

Table 4.1: Genotype Frequency to Genotype Number

Allele Frequency	Size of Population	Alleles per Person	Number of Alleles
p = 60% q = 40%	10,000 10,000	2 2	12,000 8,000
			(total 20,000)

Table 4.2: Allele Frequency to Allele Number

Interpretation of the Information

If you see this	It means this	If you see this	It means this
Carrier State (genotype)	2pq	Affected (genotype)	q2
Normal (genotype)	p2	Disease Prevalence	q2
Normal Allele	p	Diseased Allele	q
Rare Deadly Disease (allele frequency)	p=1, q=tiny fraction	Rare Deadly Disease Carrier State (geno)	2pq = carrier q2 = affected
X-Linked Recessive Prevalence	Allele freq = q Genotype freq = q		

Table 4.3

© OnlineMedEd. http://www.onlinemeded.org